



우리가 살아가는 지구를 잘
이해하기 위한 나사 (NASA)의
빅데이터(Big Data):

나사가 제공하는
초미세먼지 관측과 모의

이희교
데이터 과학자 (Data Scientist)



Jet Propulsion Laboratory
California Institute of Technology

개요

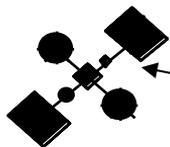
- 나사/제트추진연구소에서 다루는 빅데이터
- 미국 항공 우주국(NASA)의 지구과학 연구
- NASA가 제공하는 초미세먼지(PM_{2.5}) 정보: 위성관측과 수치모의
- 미세먼지 연구의 현재와 미래
 - 사례연구: 잘못된 해석
 - 나사와 국립 환경과학원이 미래에 제공할 자료들
- 요약

제트추진 연구소에서 다루는 빅데이터 연구들

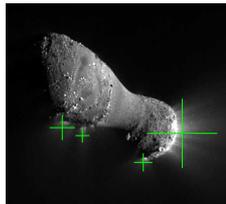
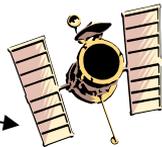


지구주변 위성에
탑재된 측정기기들

중계 위성



우주선/
화성/달



[Image from http://dailytrojan.com/2012/12/04/students-can-help-reduce-pollution/](http://dailytrojan.com/2012/12/04/students-can-help-reduce-pollution/)

일반 대중



데이터 획득과
명령 송신

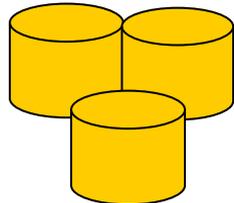
자료 가공



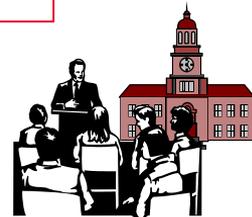
관측과 모델 자료들

(수백 페타 바이트)

1PB = 1,000 TB = 1,000,000 GB



빅데이터 분석



대학교

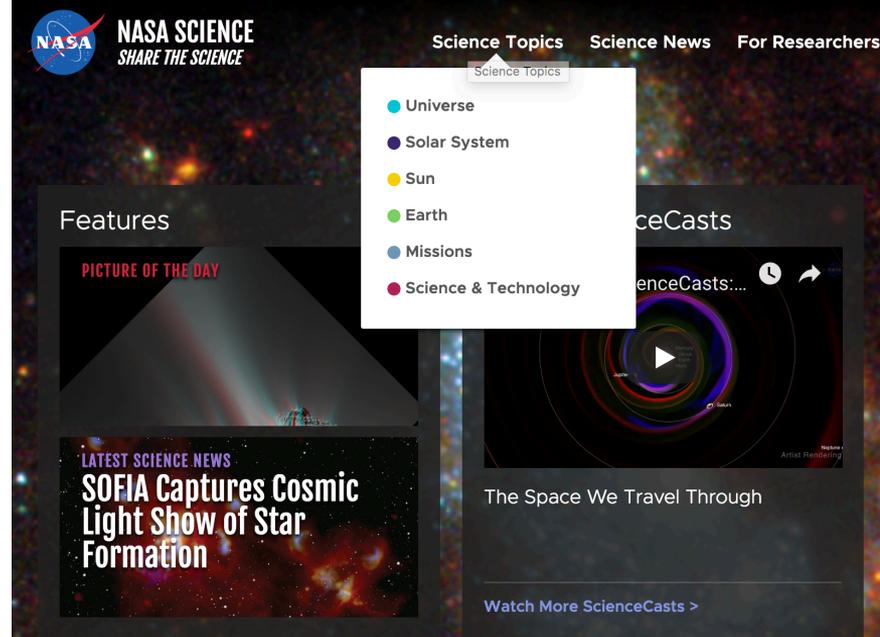
슈퍼컴퓨터를
활용한
수치모델



과학자들

나사의 지구과학

- <https://science.nasa.gov>
- 네 개의 분야
 - 지구과학
 - 천문학
 - 행성과학
 - 헬리오물리학
- 나사가 답하려고 하는 기본적인 질문들
 - 전 세계의 지구 시스템은 어떻게 변하고 있는가?
 - 이러한 변화를 일으키는 원인은 무엇인가?
 - 미래에 지구 시스템은 어떻게 변할 것인가?
 - 지구과학 연구가 어떻게 사회에 공헌할 수 있나?



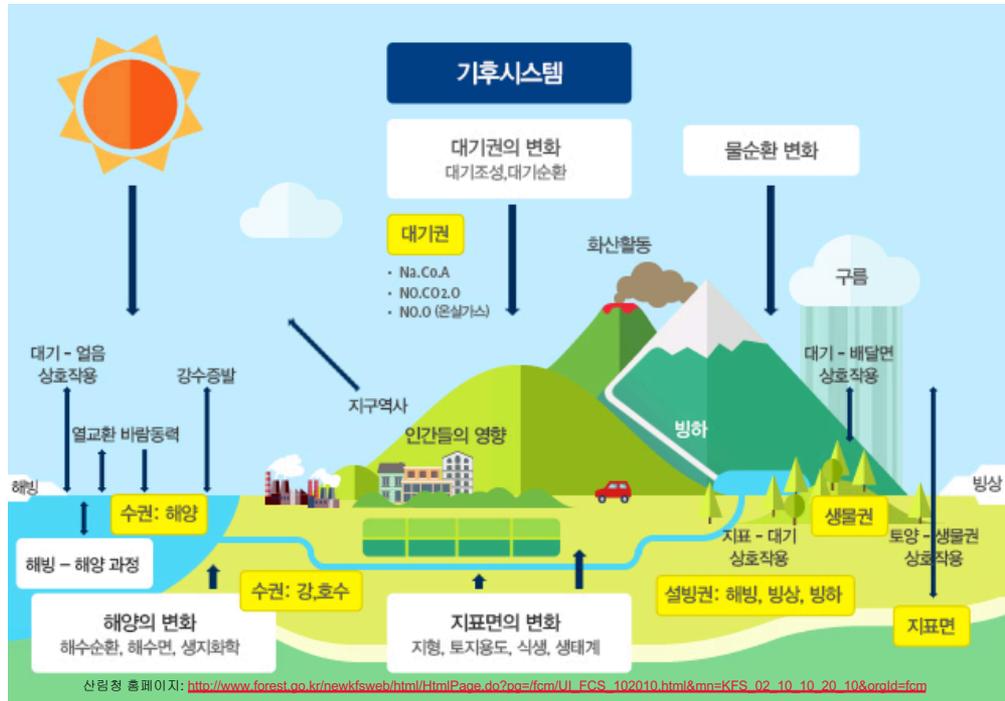
지구의 기후 시스템

- 기후시스템의 구성요소:

- 대기권
- 수권
- 생물권
- 설빙권

- 기후시스템의 동력:

- 자연적 - 태양광, 구름, 해양순환의 변화, 화산 폭발, 먼지 폭풍
- 인공적 - 온실가스, 오존, 지표면 상태변화, 미세먼지



나사의 지구과학자들이 하는 연구

- 나사의 지구과학자들은 관측을 통해 지구상의 대기와 지표면, 해양에서 일어나는 현상들을 보다 잘 이해하고 수치모델을 활용한 미래의 변화를 보다 잘 예측하려고 합니다.

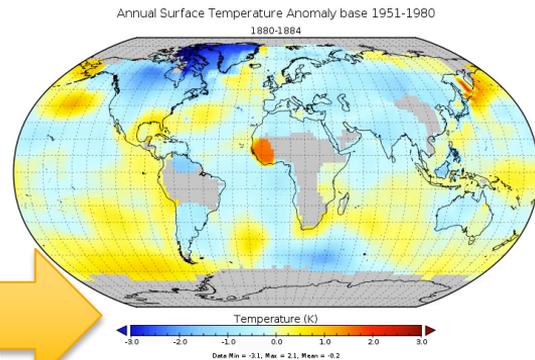
인공위성, 비행기,
지상 관측



자료 분석

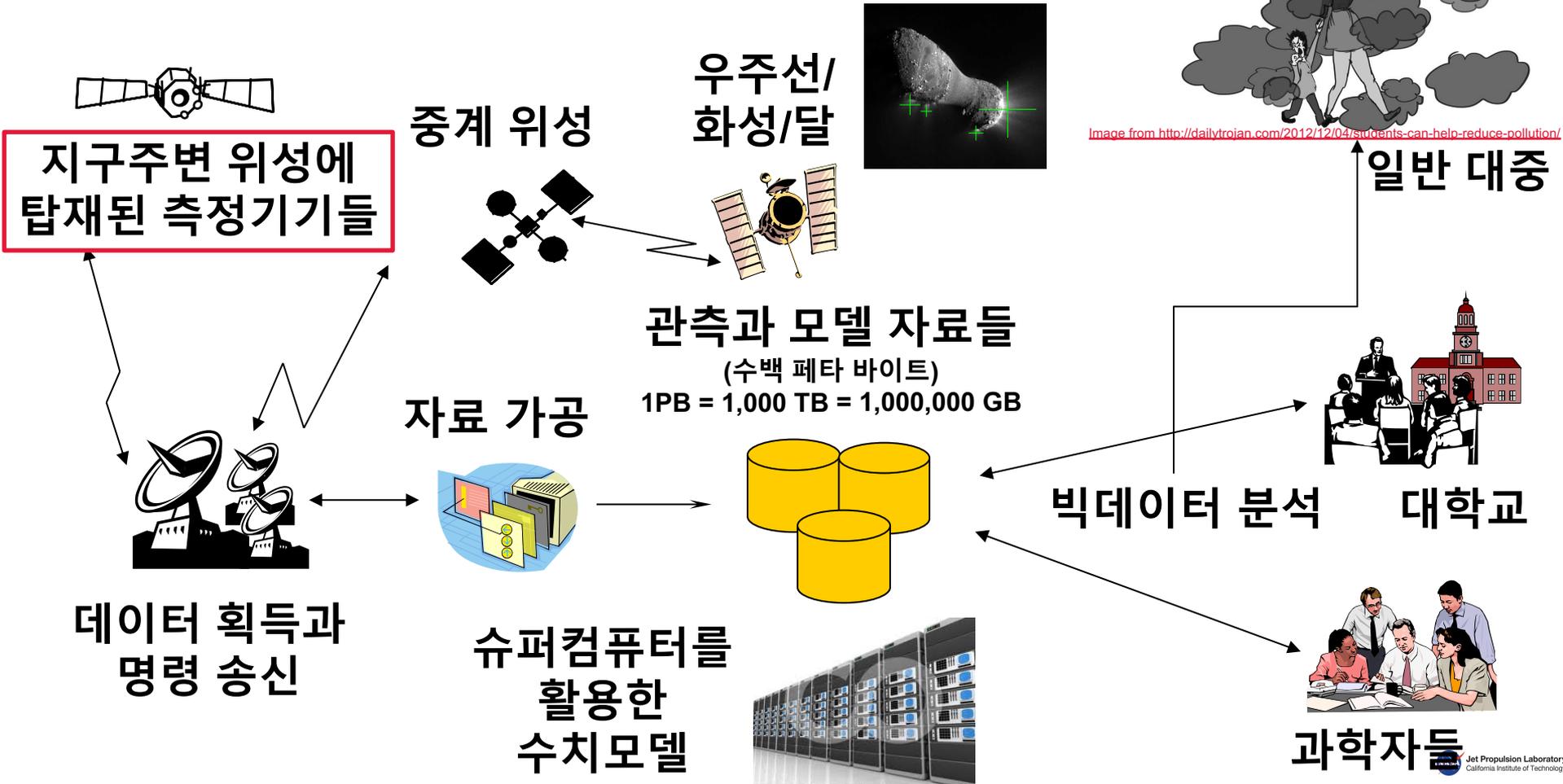


수치모델을
활용한 미래 예측



from NASA GISS
(<http://data.giss.nasa.gov/gistemp/animations/>)

제트추진 연구소에서 다루는 빅데이터 연구들



Earth Science Missions

ISS Instruments

CATS, LIS, SAGE III
 TSI-1, OCO-3, ECOSTRESS, GEDI
 CLARREO-PF

JPSS-2 Instruments

RBI, OMPS-Limb

- Formulation
- Implementation
- Primary Ops
- Extended Ops



주의: 최종 결정을
 기다리고 있는
 미션들이 있으므로
 토론용으로만
 사용가능

• 대기과학:

미세먼지, 온도,
 수증기, 바람,
 대기의 화학 조성,
 구름, 번개

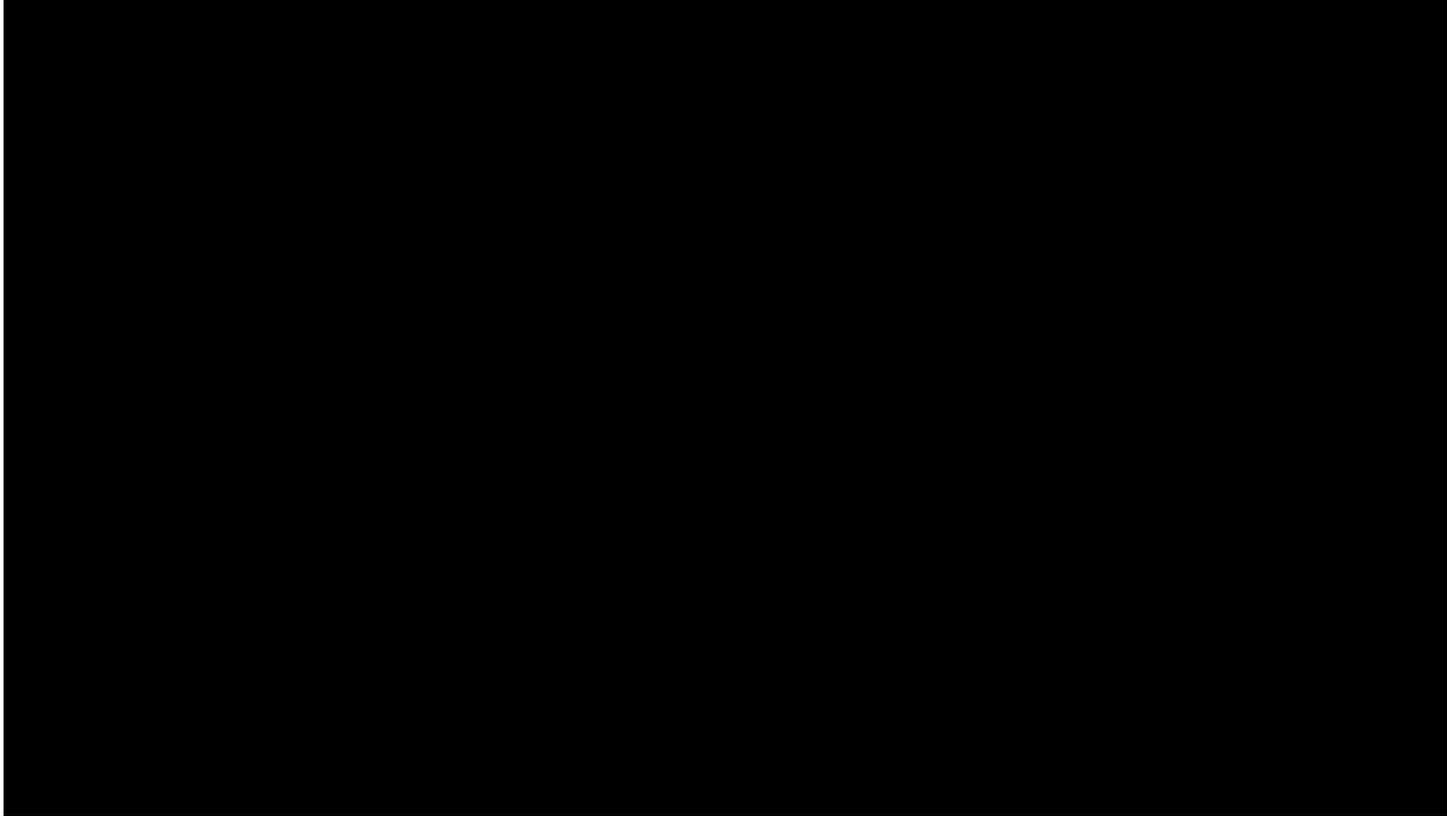
• 물과 탄소 순환:
 비, 지하수, 토양
 수분, 이산화탄소,
 식생

• 해양학:

해수면 온도,
 해수면 높이, 염분,
 바람, 극지방의
 빙하

나사의 테라(Terra) 위성: 2000년 3월 ~ 현재

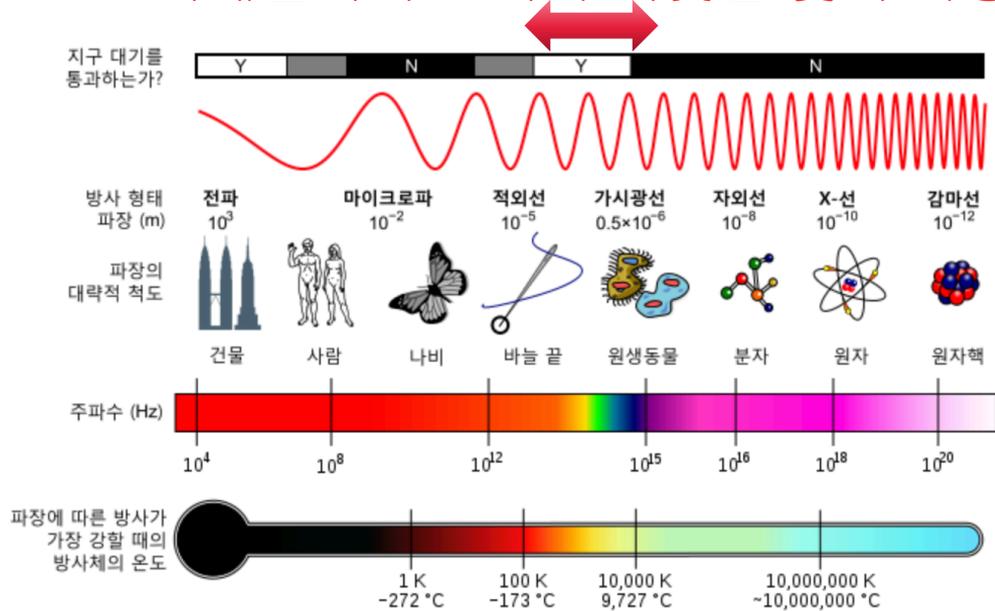
<https://youtu.be/C2uyjRGwwOs>



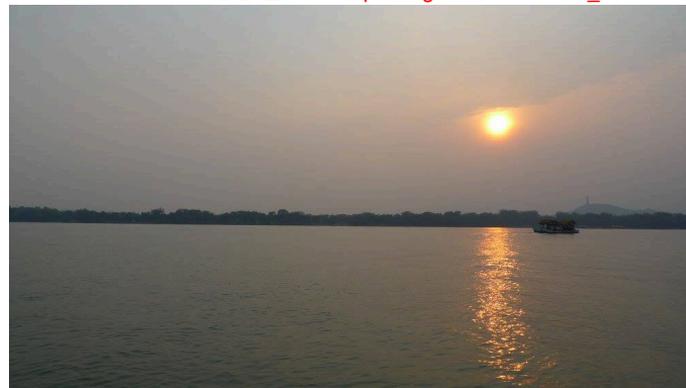
초미세먼지 (PM_{2.5})



미세먼지의 크기와 비슷한 빛의 파장



기상청 블로그: http://blog.naver.com/kma_131/221079476612



원격탐사

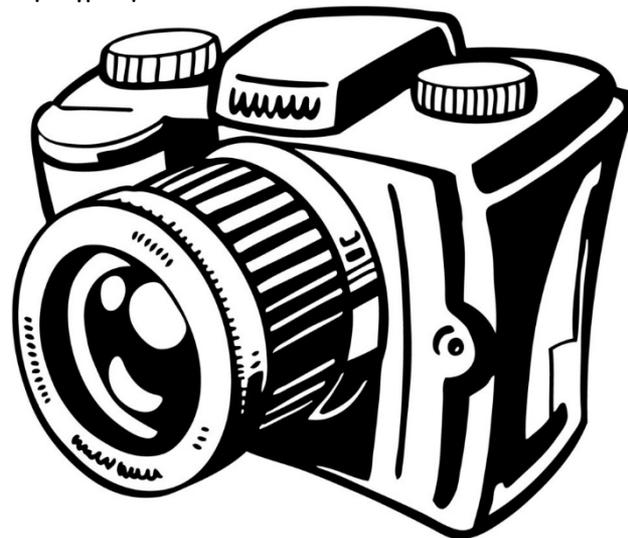
- “원격탐사는 비행기나 위성이 멀리 떨어져있는 물체나 지역으로부터 정보를 얻어내는 과학분야입니다.”

(<https://oceanservice.noaa.gov/facts/remotesensing.html>)

레이저 온도계 (<https://sciencing.com/laser-thermometers-work-4962575.html>)

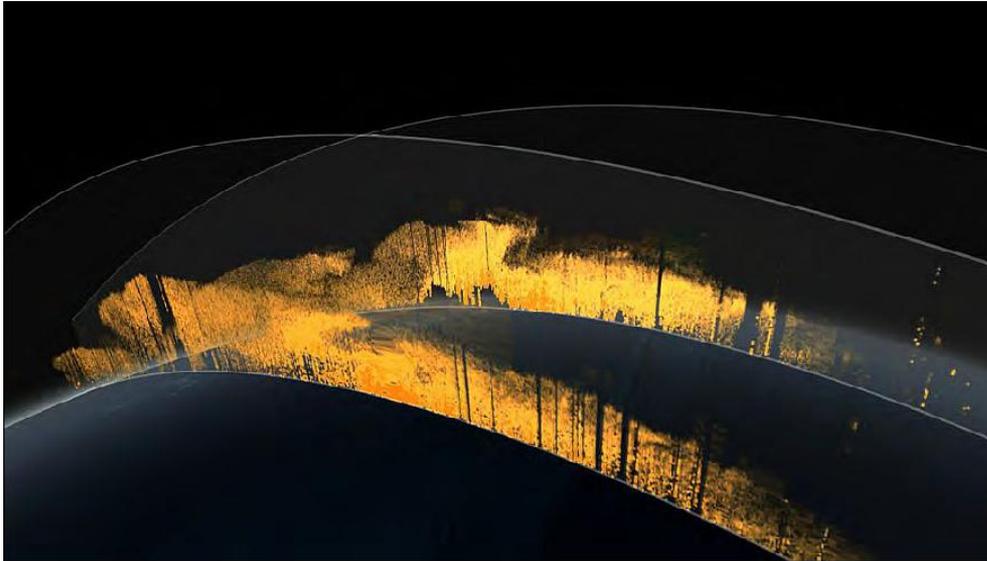


카메라



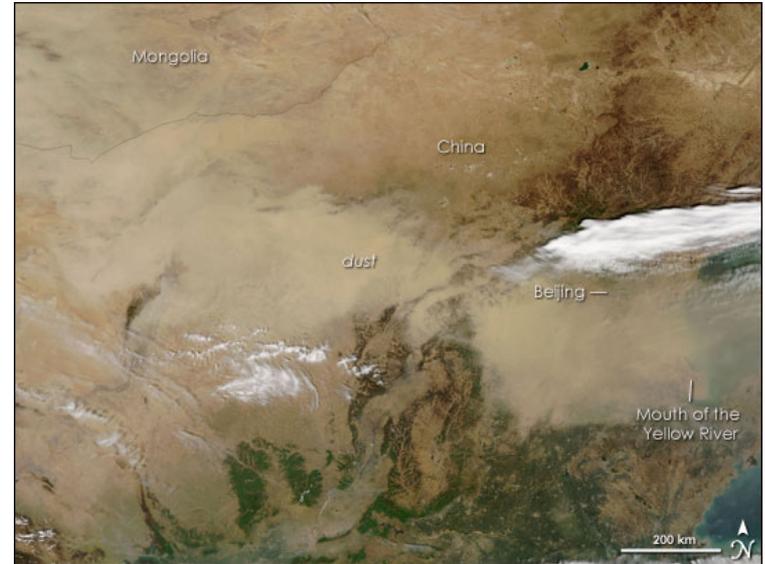
미세먼지 원격탐사: 레이저나 사진

- 레이저



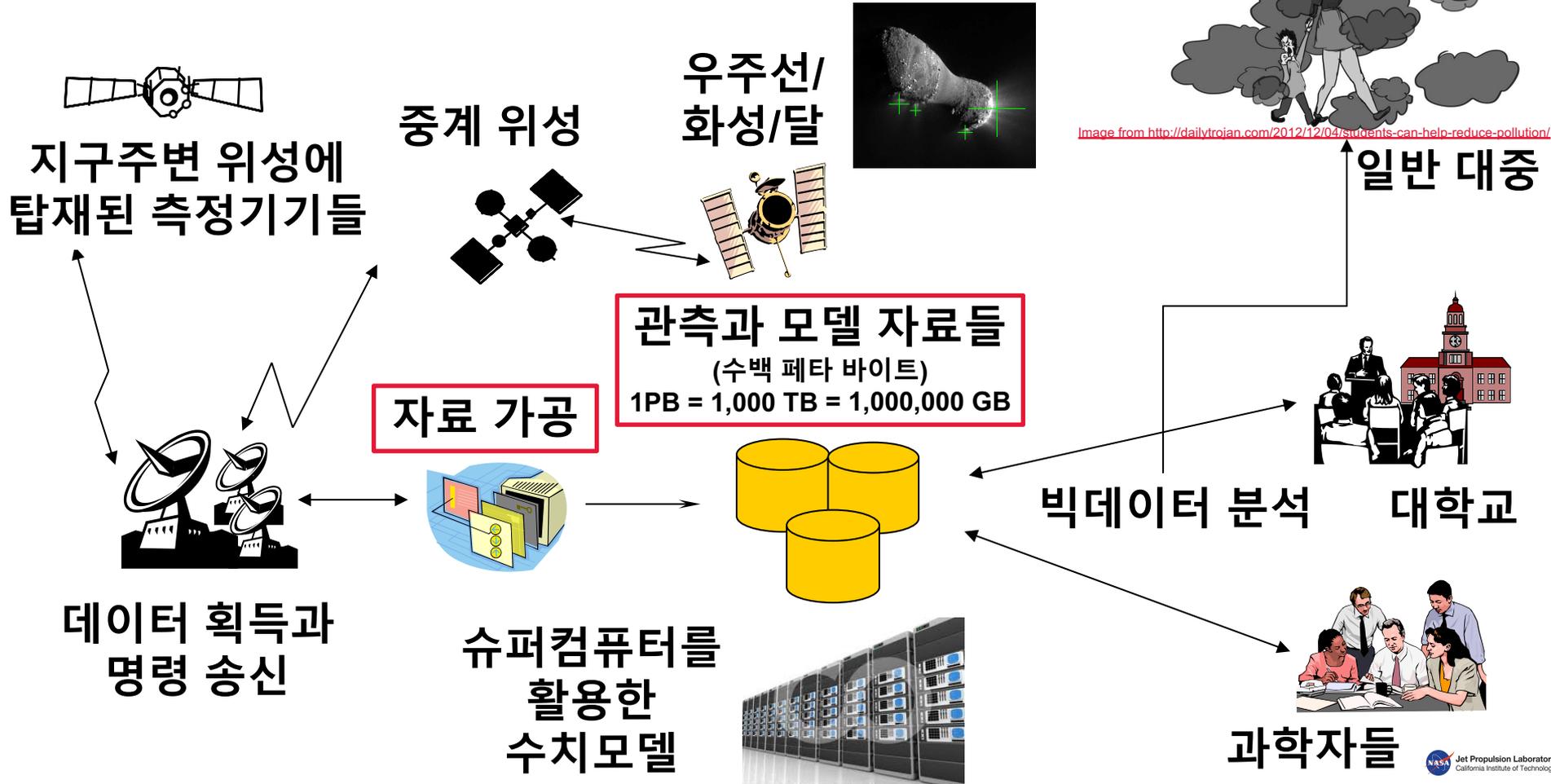
images from (<https://events.eoportal.org/web/eoportal/events/event-details/-/article/calipso>)

- 사진

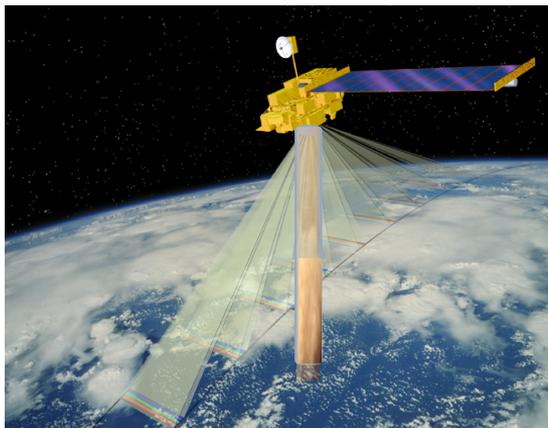


images from (<https://earthobservatory.nasa.gov/NaturalHazards/view.php?id=14861>)

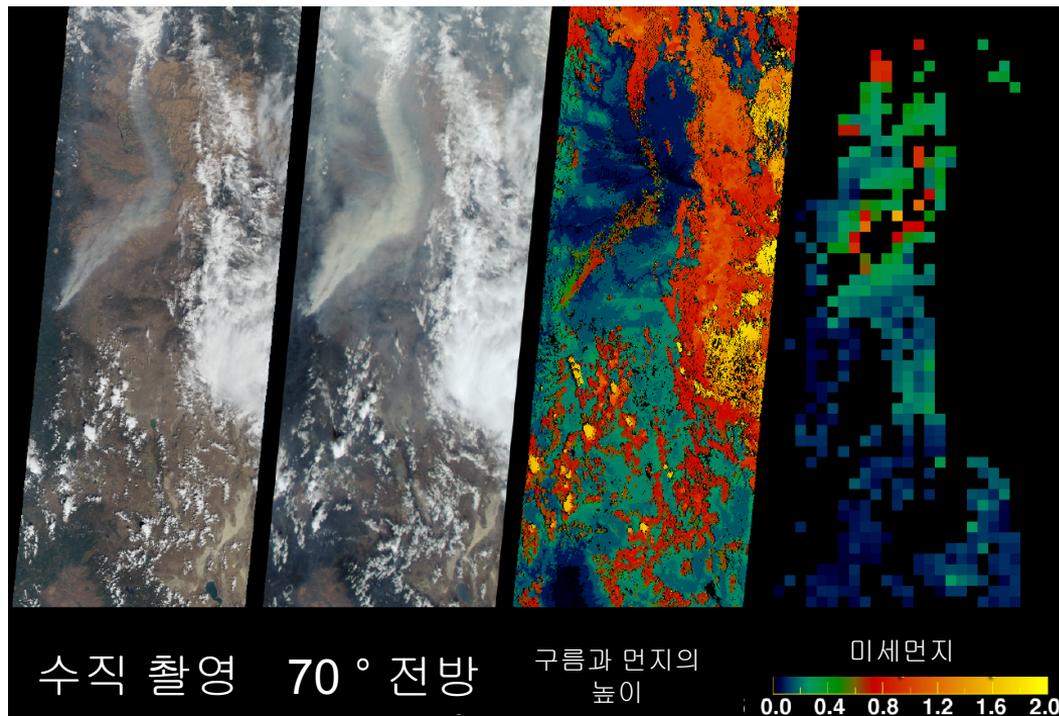
제트추진 연구소에서 다루는 빅데이터 연구들



위성자료의 가공: 위성사진 → 미세먼지

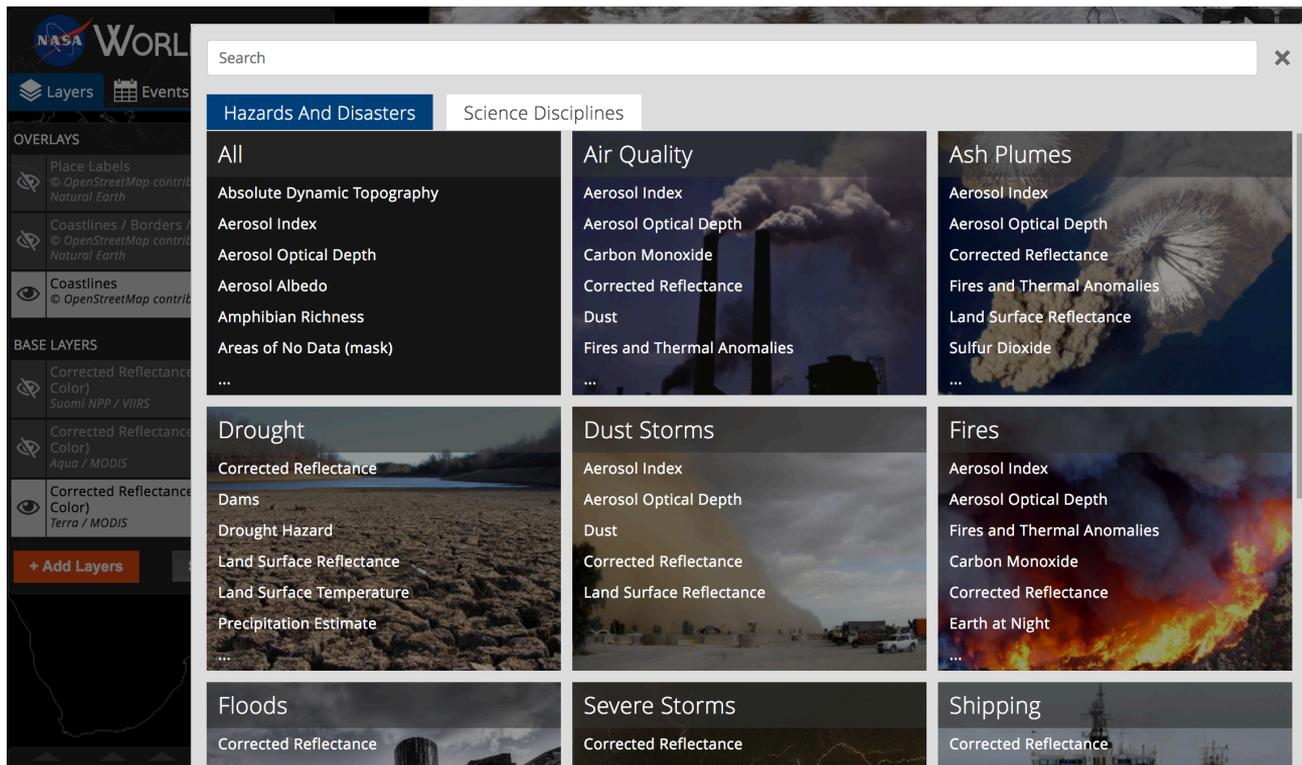


- 구름이 없어야 미세먼지가 관측가능
- 대기중의 미세먼지 총량(≠지상에서의 미세먼지 농도)만 추측가능



NASA worldview: 가공된 나사의 위성자료

(<https://worldview.earthdata.nasa.gov>)



지상에서 관측된 초미세먼지 (에어코리아, OpenAQ)



에어코리아 | 실시간자료조회 | 대기정보예보/경보 | 통계

실시간 대기정보 | 내일의 대기정보 | 시도별 대기/경보 정보 | 실시간 공간분포 | 대기오염통계

통합대기환경지수(CAI)

초미세먼지 (PM_{2.5})

미세먼지 (PM₁₀)

오존 (O₃)

이산화질소 (NO₂)

일산화탄소 (CO)

아황산가스 (SO₂)

측정망구분

도시대기 국가배경

도로변대기 교외대기

농도범위

좋음(0 ~ 15)

보통(16 ~ 35)

나쁨(36 ~ 75)

매우나쁨(76 ~ ~)

데이터 없음

기상정보

표기인함 날씨기온

풍향풍속 강수

2019년 03월 31일 18시

Home | Data | Map | Commu...

Showing the most recent values for **PM2.5**

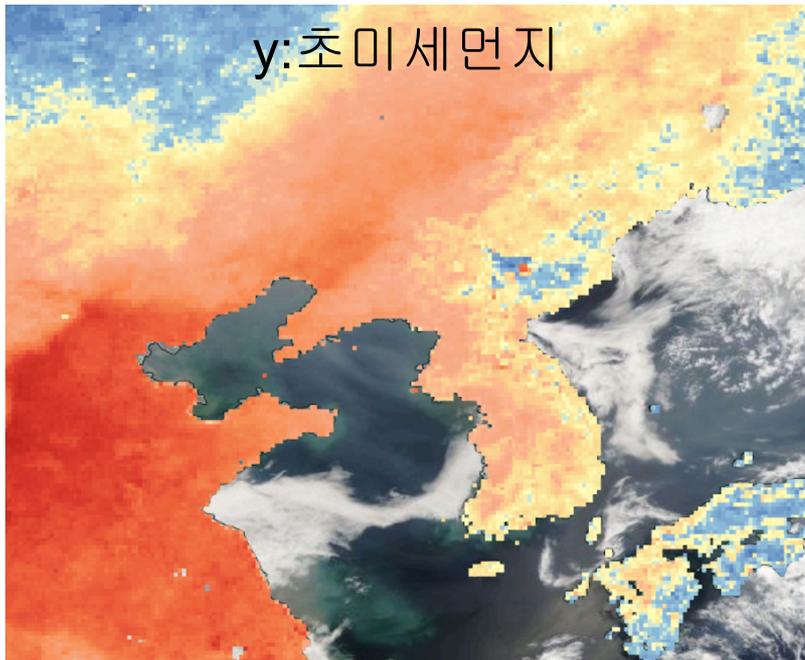
0µg/m³ | 12.2 | 24.4 | 36.7 | 48.9 | 61.1 | 73.3 | 85.6 | 97.8+

Data Disclaimer and More Information

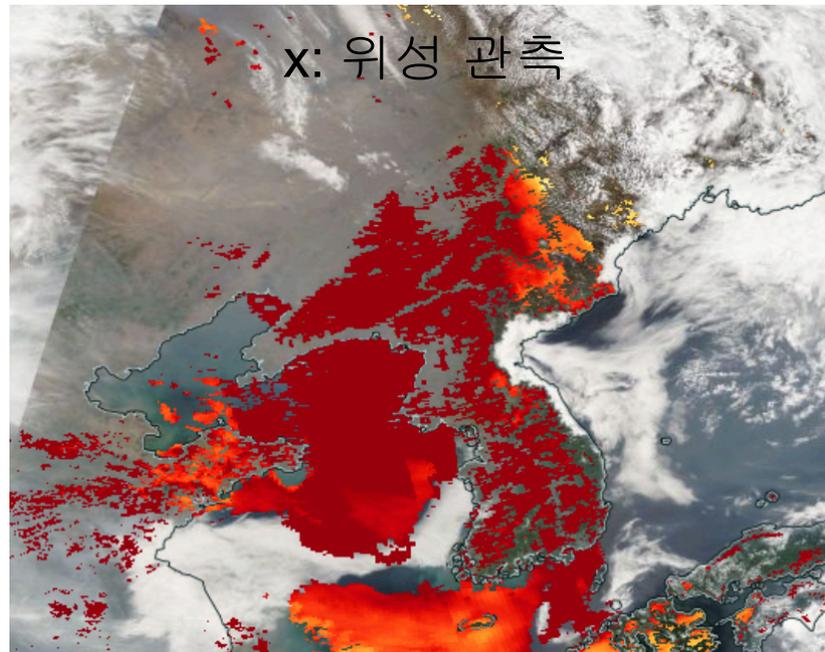
초미세먼지는 지상관측소에서만 직접 측정 가능합니다.

초미세먼지를 추측하기 위한 함수

$$y = f(x)$$



Captured on NASA Worldview



Captured on NASA Worldview

데이터 과학자들의 임무 (1): 함수를 찾아내고 업데이트

- 수 많은 위성자료중에서 어떤 자료를 써야할지를 결정
- 위성자료로부터 지상의 초미세먼지 농도를 추정하는 함수 (f)를 결정

$$\text{초미세먼지 농도} = f(\text{위성자료})$$

인공위성, 비행기,
지상 관측

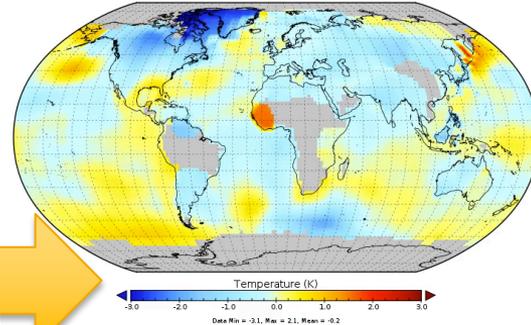


자료 분석



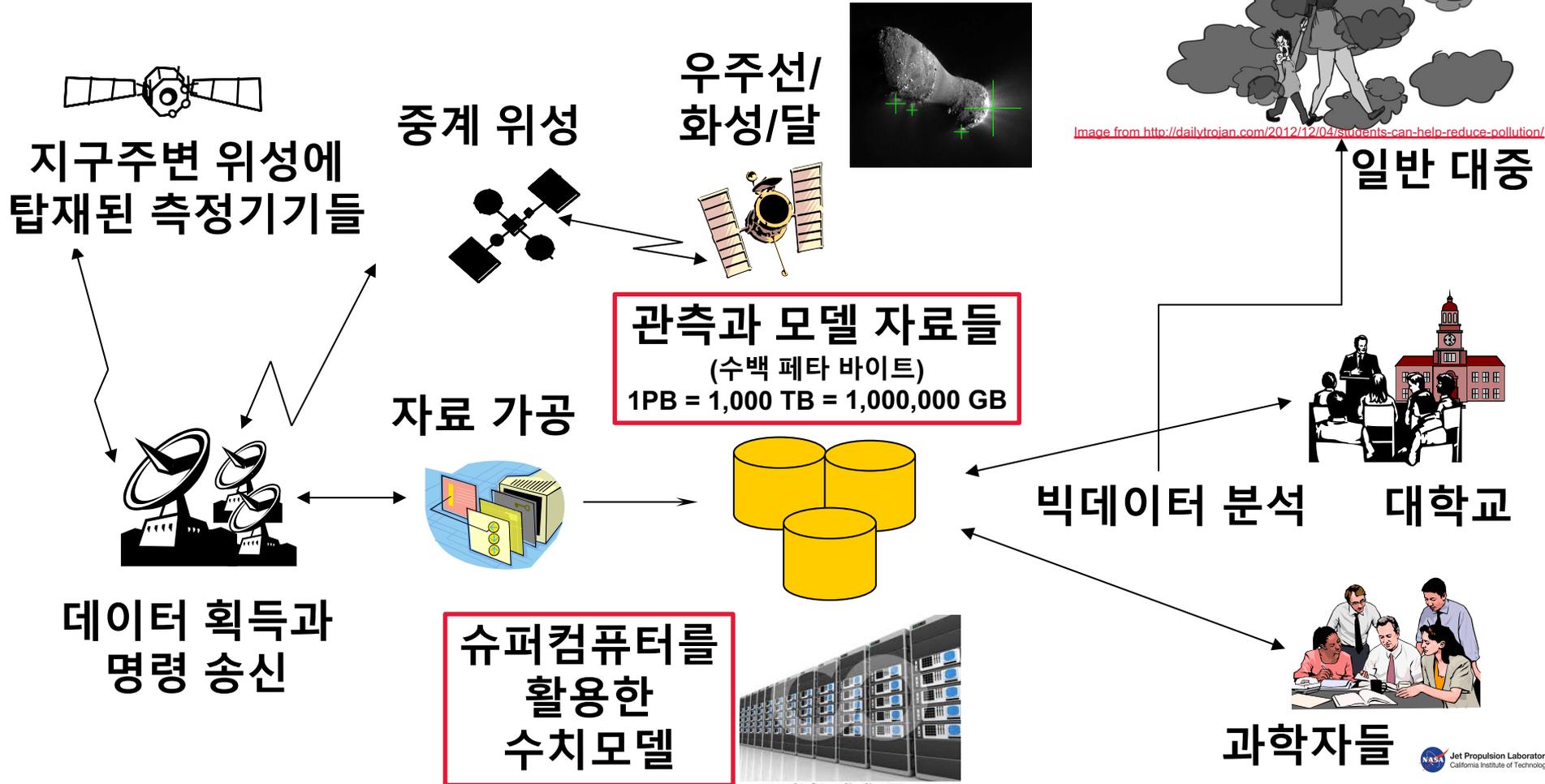
수치모델을
활용한 미래 예측

Annual Surface Temperature Anomaly base 1951-1980
1880-1884



from NASA GISS
(<http://data.giss.nasa.gov/gistemp/animations/>)

제트추진 연구소에서 다루는 빅데이터 연구들



수치모델을 이용한 미세먼지 “예측”

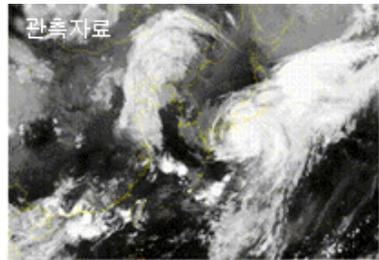
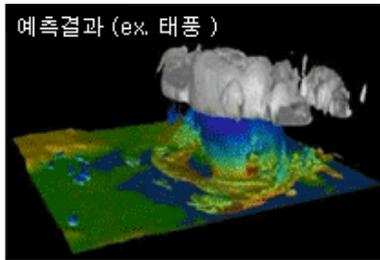
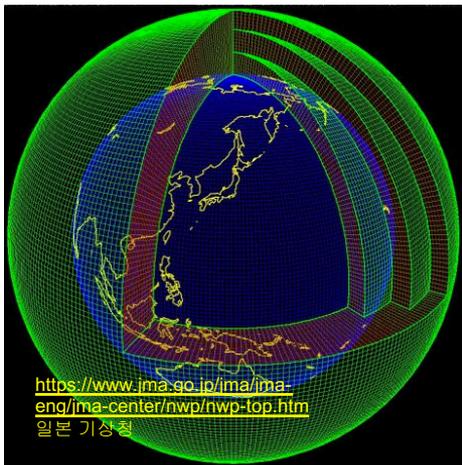
Continuity:
$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = 0$$

X - Momentum:
$$\frac{\partial(\rho u)}{\partial t} + \frac{\partial(\rho u^2)}{\partial x} + \frac{\partial(\rho uv)}{\partial y} + \frac{\partial(\rho uw)}{\partial z} = -\frac{\partial p}{\partial x} + \frac{1}{Re_r} \left[\frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} \right]$$

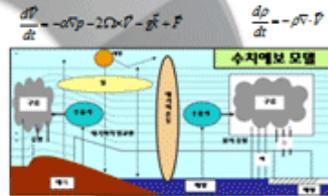
Y - Momentum:
$$\frac{\partial(\rho v)}{\partial t} + \frac{\partial(\rho uv)}{\partial x} + \frac{\partial(\rho v^2)}{\partial y} + \frac{\partial(\rho vw)}{\partial z} = -\frac{\partial p}{\partial y} + \frac{1}{Re_r} \left[\frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} + \frac{\partial \tau_{yz}}{\partial z} \right]$$

Z - Momentum:
$$\frac{\partial(\rho w)}{\partial t} + \frac{\partial(\rho uw)}{\partial x} + \frac{\partial(\rho vw)}{\partial y} + \frac{\partial(\rho w^2)}{\partial z} = -\frac{\partial p}{\partial z} + \frac{1}{Re_r} \left[\frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \tau_{zz}}{\partial z} \right]$$

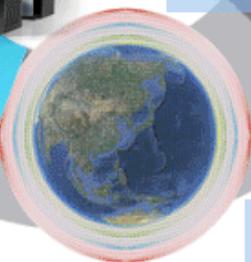
Energy:
$$\frac{\partial(E_T)}{\partial t} + \frac{\partial(uE_T)}{\partial x} + \frac{\partial(vE_T)}{\partial y} + \frac{\partial(wE_T)}{\partial z} = -\frac{\partial(u_p)}{\partial x} - \frac{\partial(v_p)}{\partial y} - \frac{\partial(w_p)}{\partial z} - \frac{1}{Re_r Pr_r} \left[\frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} + \frac{\partial q_z}{\partial z} \right] + \frac{1}{Re_r} \left[\frac{\partial}{\partial x} (u \tau_{xx} + v \tau_{xy} + w \tau_{xz}) + \frac{\partial}{\partial y} (u \tau_{xy} + v \tau_{yy} + w \tau_{yz}) + \frac{\partial}{\partial z} (u \tau_{xz} + v \tau_{yz} + w \tau_{zz}) \right]$$



• 수치모델을 단 시간에 계산하는 슈퍼컴퓨터



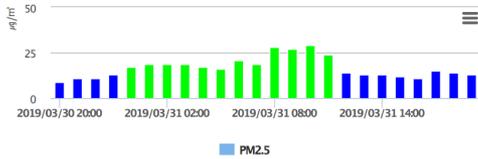
• 대기의 상태와 움직임을 재현하는 방정식계



• 방정식을 계산하기 위한 슈퍼컴퓨터용 프로그램

http://web.kma.go.kr/aboutkma/intro/supercom/model/model_concept.jsp 기상청

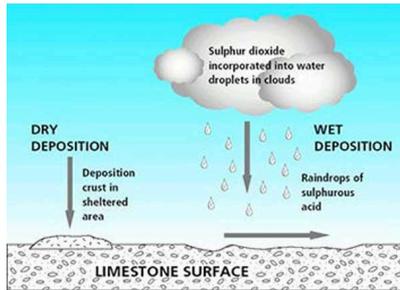
초미세먼지의 변화를 예측하는 방정식



$$(\text{초미세먼지 농도의 변화}) = (\text{수송된 초미세먼지}) + (\text{초미세먼지 석임})$$

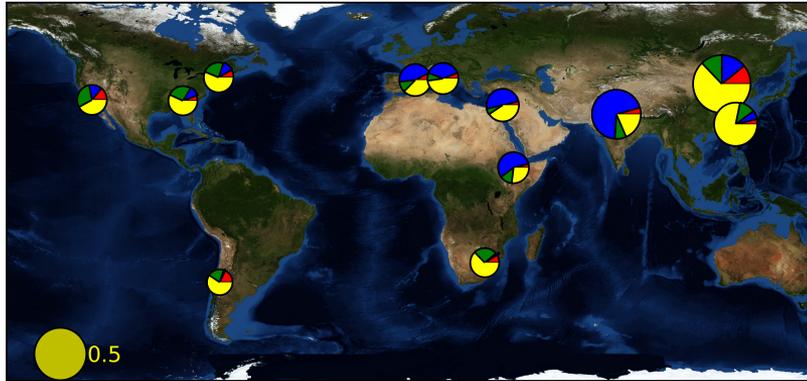


$$+ (\text{배출된 초미세먼지}) + (\text{초미세먼지의 화학반응})$$

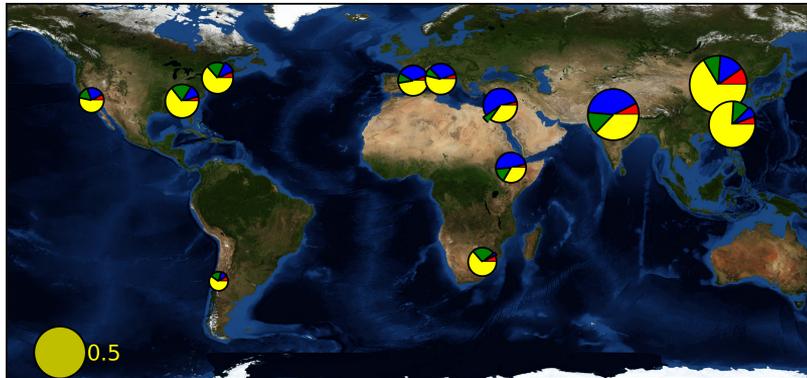


$$- (\text{낙하하는 초미세먼지}) - (\text{씻겨진 초미세먼지})$$

수치모델이 제공하는 초미세먼지의 불확실성



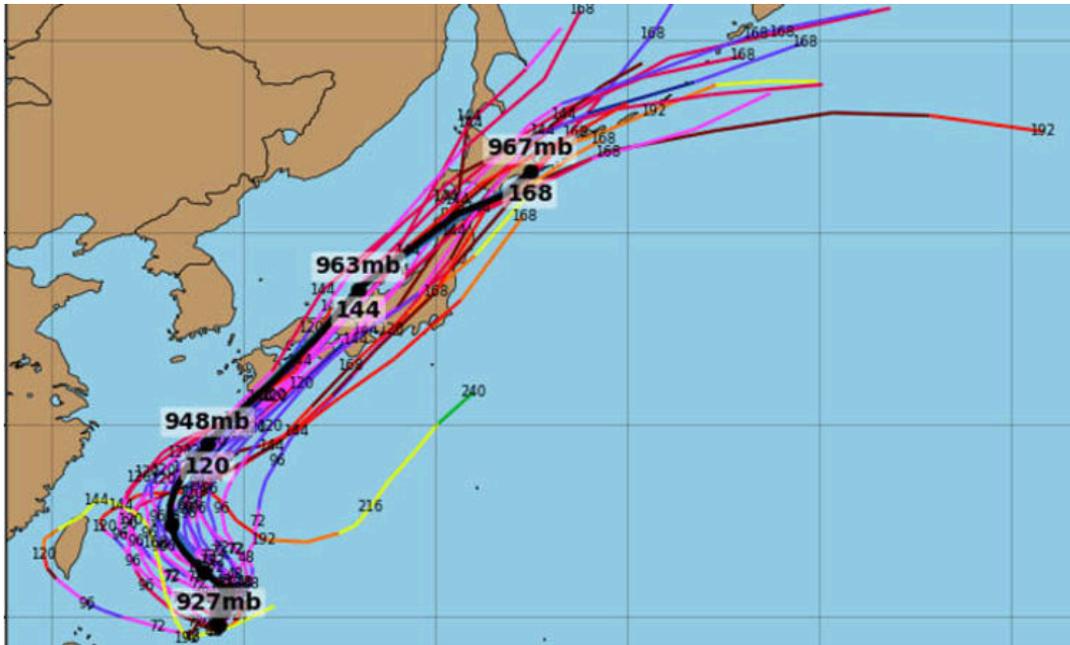
■ 숯검댕 ■ 유기탄소 ■ 먼지 ■ 황산염



- 같은 모델(나사의 GEOS-5)로 계산한 초미세먼지의 화학구성 성분입니다.
- 차이점
 - 인도 뉴델리에서의 초미세먼지 화학조성
- 가설
 - 먼지의 양을 계산하는 방법이 다를 것이다

수치모델 결과는 항상 100% 확실하지 않다!

<https://www.express.co.uk/news/world/1022439/hurricane-tracker-update-tropical-storm-leslie-update-florence-noaa-latest>



- 태풍의 경로를 예측하는 수치 모델 결과의 다양성
- 날씨를 예측하는 기상청의 수치모델과 미세먼지를 예보하는 수치모델은 기본적으로 동일
- 불확실성을 고려하여 여러가지 다른 조건에서 계산된 수치모델 자료들을 잘 취합하고 관측자료와 함께 해석하는게 중요

데이터 과학자들의 임무 (2): 수치모델 결과를 관측자료로 검증

- 여러가지 다른 조건에서 계산된 수치모델 자료들을 잘 취합하고 관측자료와 비교하는게 중요
- 관측자료와 수치모델의 차이를 바탕으로 예측의 불확실성을 파악하고 수치모델의 방정식을 개선해야한다.

인공위성, 비행기,
지상 관측

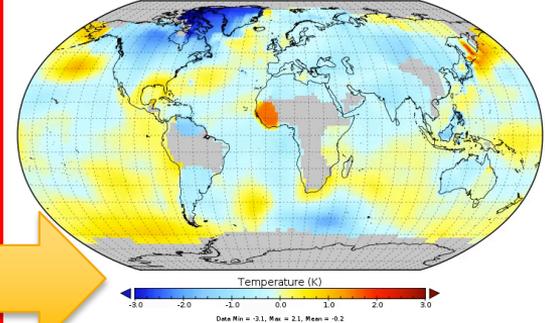


자료 분석



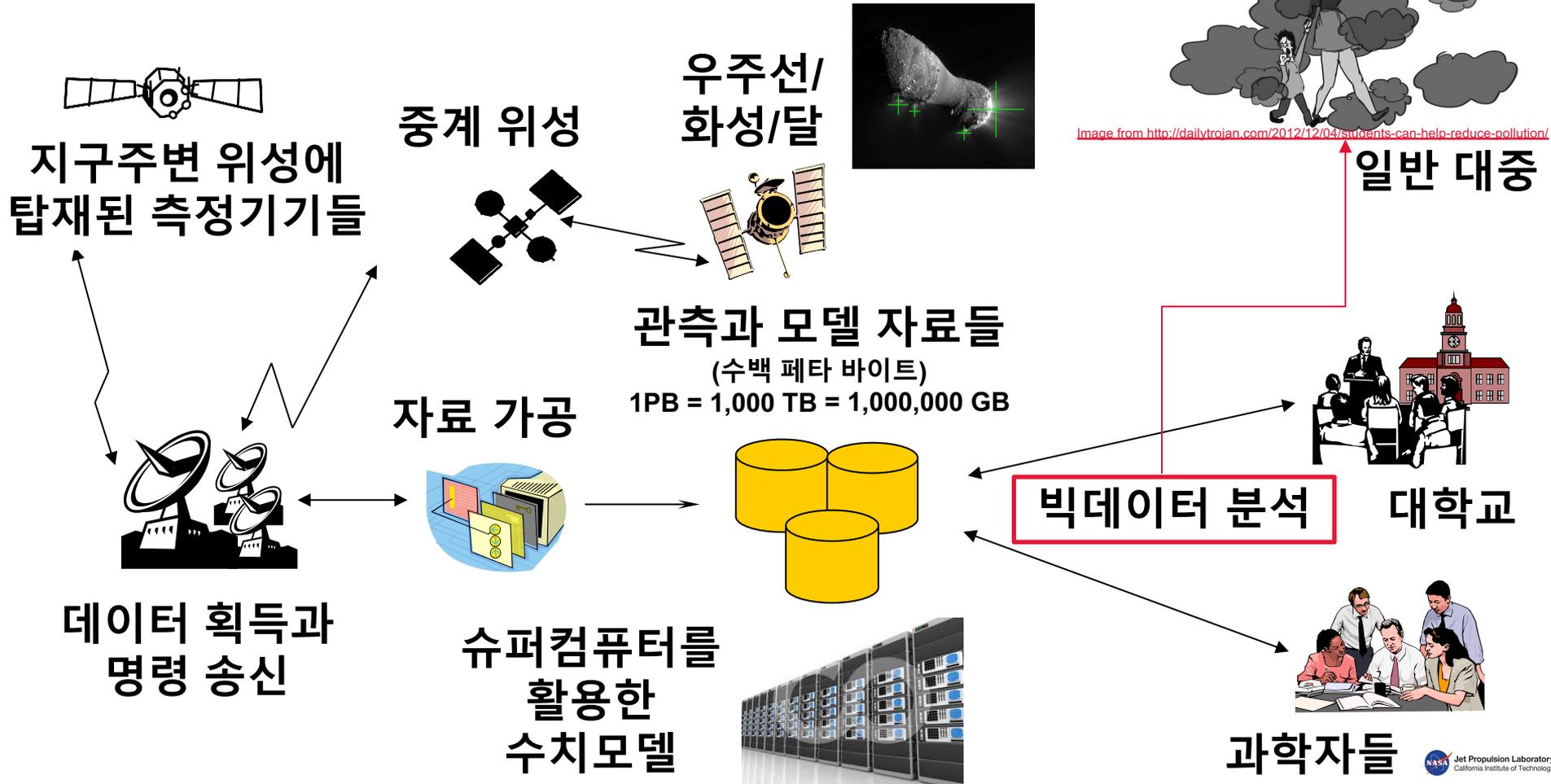
수치모델을
활용한 미래 예측

Annual Surface Temperature Anomaly base 1951-1980
1880-1884



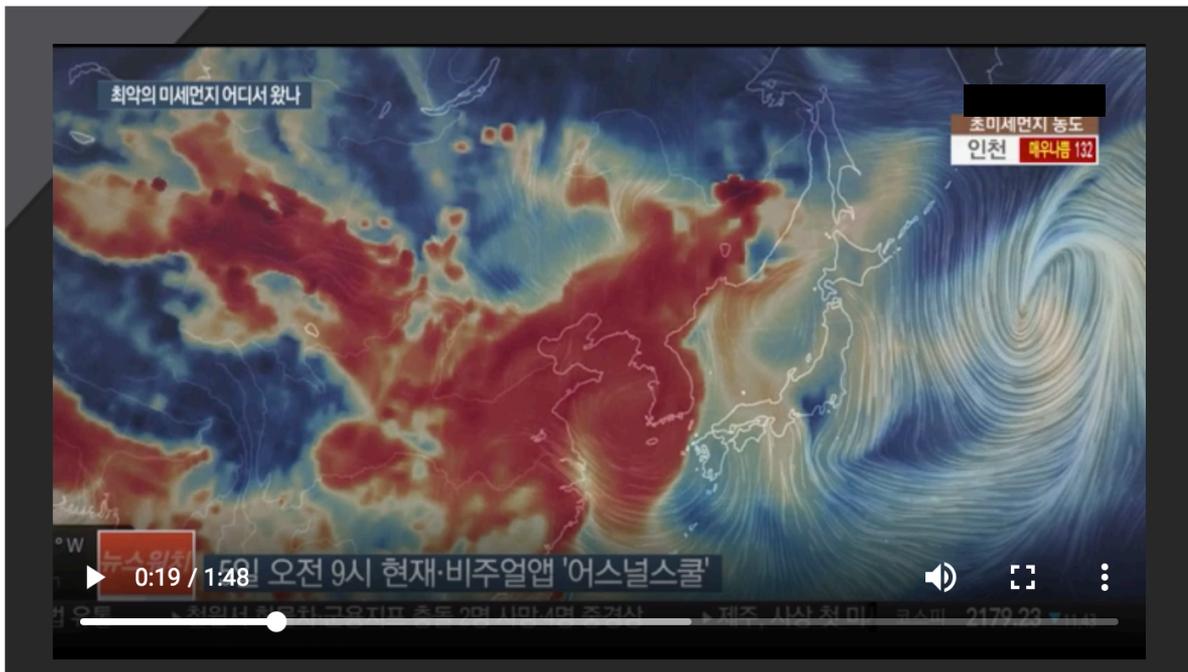
from NASA GISS
(<http://data.giss.nasa.gov/gistemp/animations/>)

제트추진 연구소에서 다루는 빅데이터 연구들





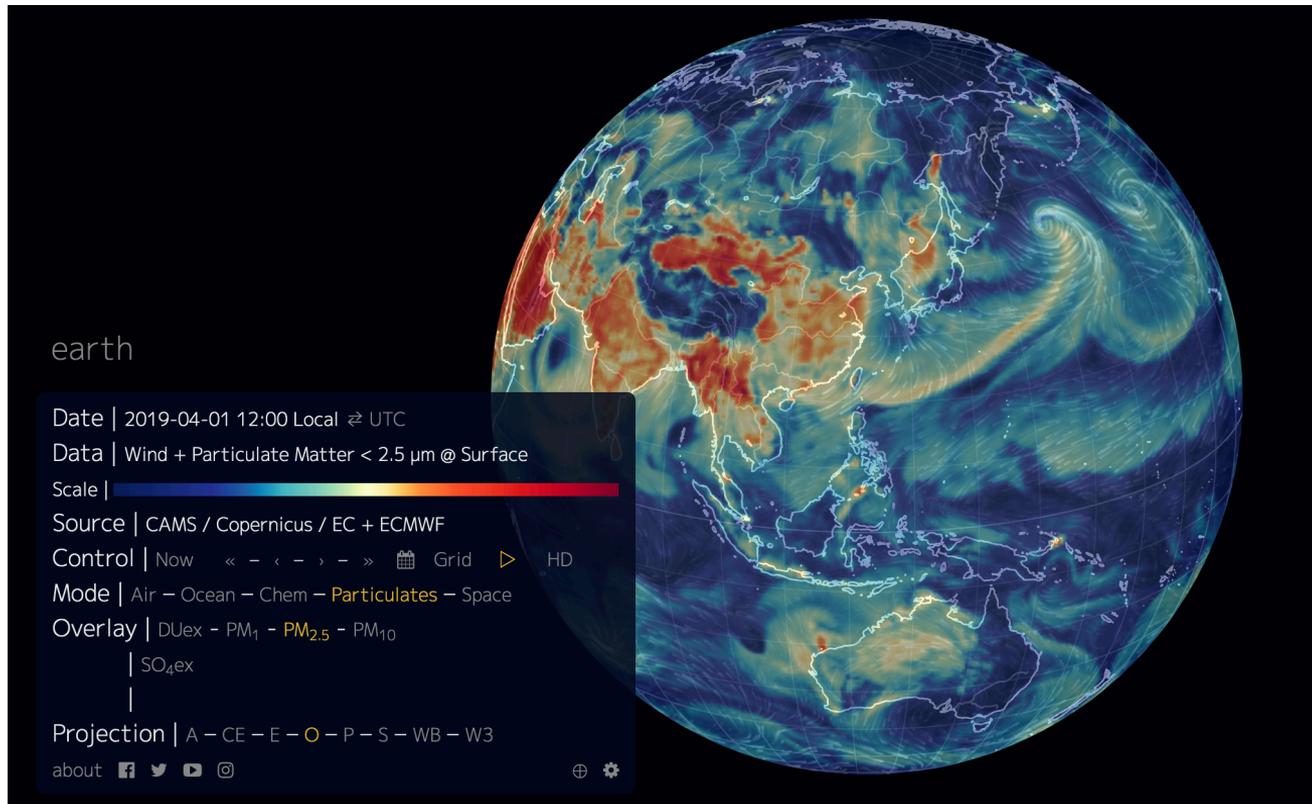
‘젯빛 한반도’... 최악 · 최장의 미세먼지 이유는?



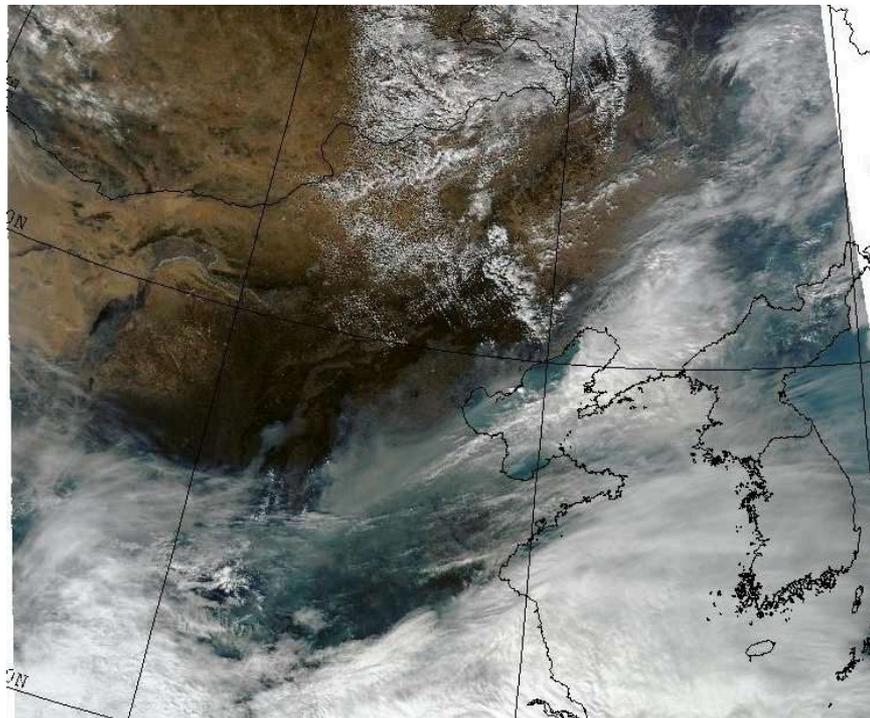
“미국 항공우주국, 나사(NASA)의 자료 등을 토대로 세계 기상정보를 시각화해 보여주는 ‘어스널스쿨’ 지도입니다.”

“어스널스쿨 (<https://earth.nullschool.net/>)”

- 나사의 슈퍼컴퓨터로 예측한 초미세먼지
- 과학적인 연구를 위해선 반드시 지상에서 관측한 초미세먼지와 비교/검증이 되어야함
- 초미세먼지가 아주 높게 나타나는 상황은 예보할 수 없음.



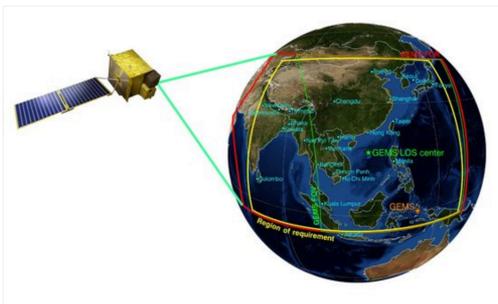
“NASA 인공위성에 딱 걸린 중국발 미세먼지”



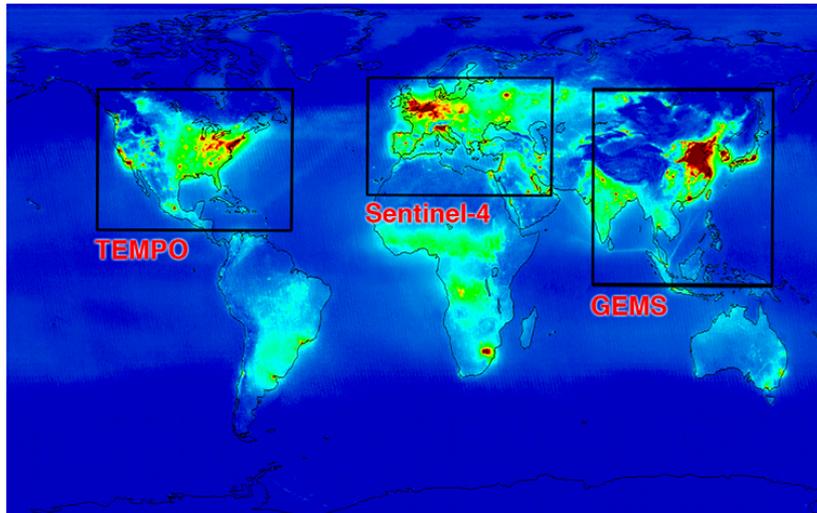
<https://www.msn.com/ko-kr/news/politics/nasa-%EC%9D%B8%EA%B3%B5-%EC%9C%84%EC%84%B1%EC%97%90-%EB%94%B1-%EA%B1%B8%EB%A6%B0-%EC%A4%91%EA%B5%AD%EB%B0%9C-%EB%AF%B8%EC%84%B8%EB%A8%BC%EC%A7%80/ar-BBv1t2>

- 가공되지 않은 위성사진
≠ 지상에서의 초미세먼지 농도
- 위성사진 → 구름이 없는 장소만
선별 → 대기중의 초미세먼지 총량
→ 함수로 지상에서의
초미세먼지를 추정 → 추정된
초미세먼지를 수치모델 &
지상관측 자료와 함께 비교
- 여러시간대에 촬영된 위성사진을
함께 활용해야 초미세먼지가
중국에서 왔다는 걸 입증 가능

“이것만 뜨면 중국발 미세먼지 손바닥 안에 있다”



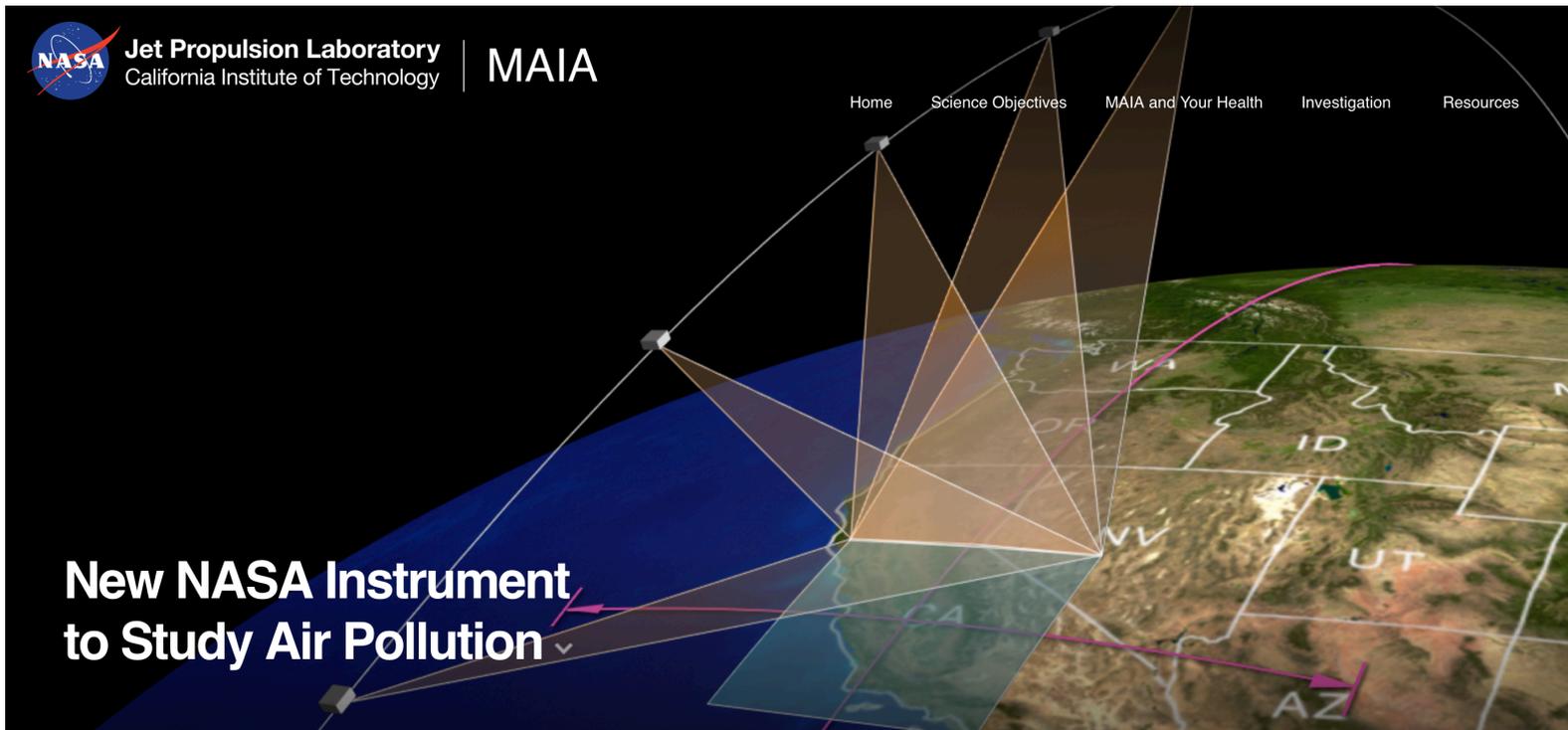
내년 발사 예정인 환경위성 쟈스의 관측 영역. 동아시아 지역 전체를 커버할 수 있다. (국립환경과학원 제공)



<http://tempo.si.edu/overview.html>

- 환경위성 쟈스 (정지궤도 복합위성 2B, 2020년 발사예정)는 초미세먼지 측정에 최적화되지 않음: 초미세먼지를 생성하는 기체 화학물질을 주로 탐지
- 기상청의 정지궤도 복합위성 2A호 (2018년 12월 발사)는 낮시간에 구름이 없는 장소에서 초미세먼지의 총량변화를 살펴보기에 보다 적합

초미세먼지 관측을 위한 나사의 새로운 위성: 마이아



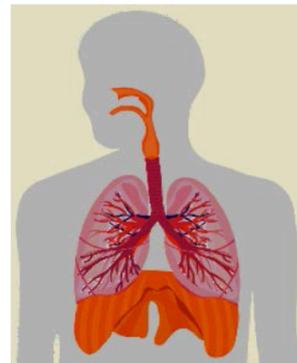
MAIA (Multi-Angle Imager for Aerosols: <https://maia.jpl.nasa.gov>)

- 전염병학 연구를 위해선 초미세먼지의 양($PM_{2.5}$)도 중요하지만 어떤 종류의 초미세먼지인지를 알아야한다

마이아 미션의 목표 (2022년 발사)



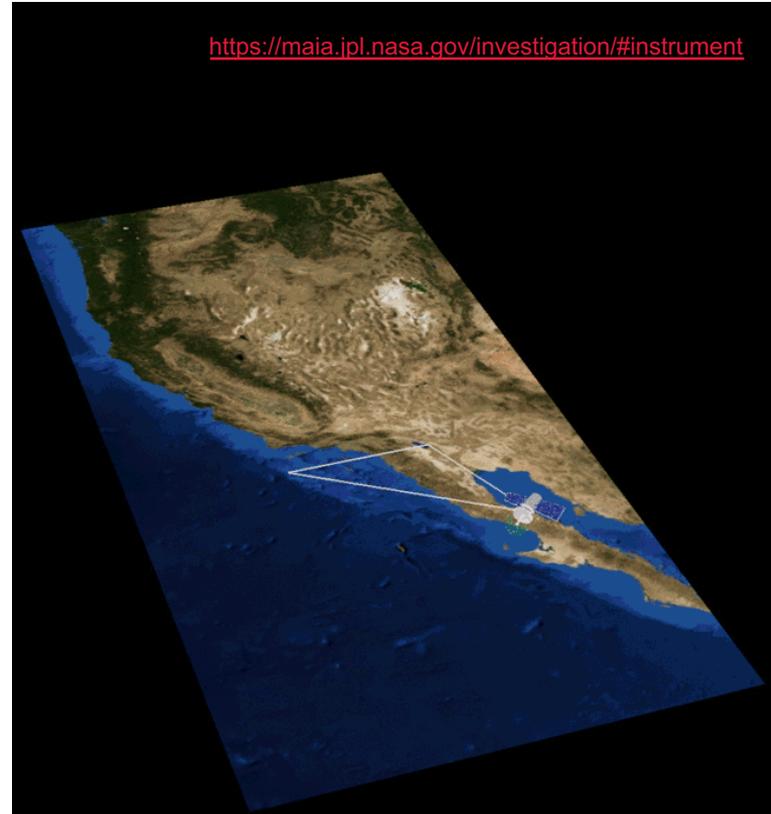
숯 검댕 먼지 황산염 등



- 어떤 종류가 초미세먼지에 단기간 동안 (며칠에서 몇주간) 노출되었을 때 가장 위험한가? 아주 높은 농도의 초미세먼지는 천식과 호흡기 질환과 관련된 입원, 정상보다 이른 죽음과 관련되어 있다.
- 임신중일 때 가장 위험한 초미세먼지는 어떤 종류인가? 높은 초미세먼지는 임신중 고혈압이나 태아의 저체중과 관련되어 있다.
- 어떤 종류의 초미세먼지가 장기간 노출시 (몇 년 이상)에 위험한가? 단기노출에 비해 장기 노출의 영향이 훨씬 큰 것으로 알려져있다. 장기적으로 고농도의 초미세먼지에 노출되면 심장병과 심장마비가 일어날 확률이 높아지며 폐암이나 다른 호흡기 질환과도 관련되어 있다.

요약(1): 빅데이터와 초미세먼지 연구

- 나사는 1) 지구궤도를 돌고 있는 수많은 위성들과 2) 슈퍼컴퓨터를 활용한 수치모델을 이용해서 초미세먼지를 비롯한 지구의 기후 시스템에 대한 정보를 제공한다.
- 나사의 차세대 마이아 위성 (2022년 발사예정) 은 미세먼지의 양 뿐만 아니라 어떤 종류의 미세먼지가 우리의 건강에 영향을 미치는지를 파악할 수 있는 중요한 단서를 제공할 것이다.



요약(2): 빅데이터와 초미세먼지 연구

- 미세먼지와 관련된 빅데이터를 다룰 땐 활용한 자료 (지상관측, 위성관측, 수치 모델)의 양이 많은 게 중요한 게 아니다. 수 많은 자료중에서 어떤 자료를 선별적으로 잘 활용하는지 배우는 것이 더 중요하다.
- 나사에서 초미세먼지를 연구하는 데이터 과학자들의 역할:
 - 위성자료를 이용해서 초미세먼지를 추측하는 함수를 찾아내고 업데이트
 - 수치모델 결과를 관측자료로 검증
- 위성사진이나 수치모델자료만으로 미세먼지에 대한 결론을 내리는 것은 과학적이지 않은 접근이다.



Jet Propulsion Laboratory
California Institute of Technology

email: huikyo.lee@jpl.nasa.gov